Marquette University
e-Publications@Marquette

Mathematics, Statistics and Computer Science Faculty Research and Publications

Mathematics, Statistics and Computer Science, Department of

9-1-2010

A Class of Models for Uncorrelated Random Variables

Nader Ebrahimi
Northern Illinois University

Gholamhossein Hamedani
Marquette University, gholamhoss.hamedani@marquette.edu

Ehsan S. Soofi
University of Wisconsin - Milwaukee

Hans Volkmer
University of Wisconsin - Milwaukee

Accepted version. *Journal of Multivariate Analysis*, Vol. 101, No. 8 (September 2010): 1859-1871. DOI.

A Class of Models for Uncorrelated Random Variables

By Nader Ebrahimi, G.G. Hamedani, Ehsan S. Soofi, and Hans Volkmer

We consider the class of multivariate distributions that gives the distribution of the sum of uncorrelated random variables by the product of their marginal distributions. This class is defined by a representation of the assumption of sub-independence, formulated previously in terms of the characteristic function and convolution, as a weaker assumption than independence for derivation of the distribution of the sum of random variables. The new representation is in terms of stochastic equivalence and the class of distributions is referred to as the summable uncorrelated marginals (SUM) distributions. The SUM distributions can be used as models for the joint distribution of uncorrelated random variables, irrespective of the strength of dependence between them. We provide a method for the construction of bivariate SUM distributions through linking any pair of identical symmetric probability density functions. We also give a formula for measuring the strength of dependence of the SUM models. A final result shows that under the condition of positive or negative orthant dependence, the SUM property implies independence.

1. Introduction

We present models for the joint distribution of uncorrelated variables that are not independent, but the distribution of their sum is given by the product of their marginal distributions. We refer to these models as the summable uncorrelated marginals (SUM) distributions. These models are developed utilizing the assumption of sub-independence which has been used previously as a weaker assumption than independence for the derivation of the distribution of the sum of random variables.

Let $\mathbf{X} = (X_1, \ldots, X_p)'$ be a random vector with probability distribution function F and characteristic function $\Psi(\mathbf{t})$. Components of $\mathbf{X}$ are said to be sub-independent if

$$\Psi(\mathbf{t}) = \prod_{i=1}^{p} \Psi_i(t), \qquad \forall \mathbf{t} = (t, \ldots, t)' \in \mathfrak{R}^p,$$

where $\Psi_i(t)$ is the characteristic function of X_i. For $p = 2$, (1) was utilized in [1] to construct bivariate models with normal marginals and Durairajan [2] referred to this assumption as sub-independence. Hamedani and Walter [3] proved several versions of the Central Limit Theorem for the sequence of random variables that satisfy (1). The assumption of sub-independence can replace that of independence in most of the theorems in probability and statistics which deal with the distribution of the sum of the random variables, rather than the joint distribution of the summands; see [4] for more references.

Independence implies (1) and the variables that satisfy (1) must be uncorrelated. A representation in terms of convolution usually accompanies (1) to provide further interpretation. In Section 2, we give an alternative representation of (1) in terms of stochastic equivalence, which can be interpreted more intuitively as the basis for the SUM models. This representation naturally leads to the mutual information (see, e.g., [5,6]) which is a measure of dependence between the variables. We provide a series expansion for the mutual information of a class of distributions which includes the Farlie–Gumbel–Morgenstern (F–G–M) family and two families of SUM distributions developed in this paper.

Numerous general methods are available for constructing a joint distribution by linking given univariate distributions as the marginals, see for example [7–14]. In Section 3, we present a method for the general construction of bivariate SUM distributions by linking univariate symmetric distributions. We show that Kendall's tau and Spearman's rho for these models are zero. However, these are not properties of all SUM models. We also provide a formula for the mutual information measure for assessing the extent of dependence of the proposed family of SUM models.

The SUM models are capable of capturing weak and strong nonlinear dependence between variables. In Section 4 we compare the strength of dependence that is captured by some bivariate SUM models with other models. The illustrations include discrete and continuous examples. We derive the mutual information formula for the F–G–M family and show that its upper bound is less than that for some SUM examples. In contrast, Kendall's tau and Spearman's rho for these examples are zero, but for the F–G–M family, in general, are not. We construct a continuous SUM family of distributions for random variables that are not independent but all their polynomial functions are uncorrelated, $\mathrm{cov}(X_1^n, X_2^m) = 0$ for all $m, n = 1,2, \ldots$. We obtain the mutual information formula for this family and compare it with the dependence measure for a non-SUM family with the same dissociation property.

Often it is of interest to identify conditions under which a weak dissociation such as uncorrelatedness is equivalent to independence. In Section 5, we discuss generalizations of (1) in the multivariate case and give a few examples. We provide a result showing that sub-independence under the well-known notions of positive and negative orthant dependence is equivalent to independence. Section 6 gives brief conclusions.

2. Representation of SUM and Mutual Information

Let F be the probability distribution function of $\mathbf{X} = (X_1, X_2)$, and $\mathbf{X}^* = (X_1^*, X_2^*)$ denote the random vector with probability distribution function $F^*(x_1, x_2) = F_1(x_1)F_2(x_2)$, where $F_i, i = 1,2$ is the marginal probability distribution function of X_i.

Definition 1. F is said to be a summable uncorrelated marginals (SUM) bivariate distribution if $X_1 + X_2 \overset{st}{=} X_1^* + X_2^*$, where $\overset{st}{=}$ denotes the stochastic equality. Random variables with a SUM joint distribution are referred to as SUM random variables.

It is clear that the SUM and sub-independence are equivalent, so the two terminologies can be used interchangeably. It is also clear that the class of SUM random variables is closed under scalar multiplication and addition under independence. That is, if $\boldsymbol{X} = (X_1, X_2)$ is a SUM random vector, so is $a\boldsymbol{X}$, and if $\boldsymbol{Y} = (Y_1, Y_2)$ is another SUM random vector independent of **X**, then **X** + **Y** is also a SUM random vector. However, the SUM property is directional in that X_1 and X_2 being SUM random variables does not imply that X_1 and aX_2 are SUM. Definition 1 can be generalized to any specific direction by $a_1X_1 + a_2X_2 \overset{st}{=} a_1X_1^* + a_2X_2^*$.

The discrepancy between F and F^* is only due to the dependence between X_1 and X_2, thus any discrepancy function between these two distributions is a measure of dependence. Kullback–Leibler discrimination information between F and F^* gives the mutual information between X_1 and X_2:

$$M(X_1, X_2) = K(F\colon F^*) = \int\int_{\Re^2} \log \frac{f(x_1,x_2)}{f^*(x_1,x_2)} \mathrm{d}F(x_1, x_2) \geq 0, \tag{2}$$

where $\mathrm{d}F(x_1, x_2) = f(x_1, x_2)\mathrm{d}x_1\mathrm{d}x_2$ for continuous and $\mathrm{d}F(x_1, x_2) = f(x_1, x_2)$ for discrete variables, and $f^*(x_1, x_2) = f_1(x_1)f_2(x_2)$, provided that $F(x_1, x_2)$ is absolutely continuous with respect to the reference distribution $F^*(x_1, x_2)$. The equality in (2) holds if and only if $f(x_1, x_2) = f^*(x_1, x_2)$ almost everywhere; i.e., if and only if X_1 and X_2 are independent. Other representations of the mutual information are:

$$\begin{aligned} M(X_1, X_2) &= H(X_1) + H(X_2) - H(X_1, X_2) \\ &= H(X_1^*, X_2^*) - H(X_1, X_2), \end{aligned} \tag{3}$$

where $H(\mathbf{X}) = -\int_{\Re^p} \log f(\mathbf{x})\mathrm{d}F(\mathbf{x}),\ p = 1,2,$ is the Shannon entropy. The second equality is due to the property that Shannon information is additive for independent random variables, and signifies that in general, f is more concentrated than f^*. For the continuous case, $M(X_1, X_2)$ is usually calibrated with the mutual information of bivariate normal distribution, $M(Y_1, Y_2) =$

$-\frac{1}{2}\log(1-\rho^2)$, where ρ is the product moment correlation coefficient of the bivariate normal model.

An important property of $M(X_1, X_2)$ is invariance under one-to-one transformations of X_i. In particular, the probability integral transformation $u_i = F(x_i)$ gives $M(X_1, X_2) = -H[c(u_1, u_2)]$, where $c(u_1, u_2)$ is the copula density of the joint distribution. This is easily seen from (3) when the distributions of $U_i, i = 1, 2$ are uniform over [0,1] and $H(U_i) = 0$.

We also use Kendall's tau τ and Spearman's rho ρ_s; see [6]. For continuous distributions,

$$\tau = 4\int\int_{\Re^2} F(x_1, x_2) f(x_1, x_2)\,\mathrm{d}x_1\mathrm{d}x_2 - 1 \tag{4}$$

$$\rho_s = 12\int\int_{\Re^2} F_1(x_1)F_2(x_2)\, f(x_1, x_2)\mathrm{d}x_1\mathrm{d}x_2 - 3. \tag{5}$$

These measures are invariant under strictly increasing transformations. However, since in general, unlike the mutual information, $\tau = 0$ and $\rho_s = 0$ do not imply independence, these measures cannot capture complicated dependence structures. For a SUM model, both measures can be nonzero, one of them can be zero while the other one is not, and both can be zero without the variables being independent. We will provide examples showing these cases.

A bivariate SUM copula is a SUM distribution on the unit square $[0, 1]^2$ with uniform marginals.

Lemma 1. *For any SUM copula,* $\rho_s = 0$*.*

Proof. This follows from the fact that for copulas $\rho_s = \rho$ (see, e.g., [8], p. 156). □

A family of SUM models with $\tau = \rho_s = 0$ will be presented in Section 3. We need the following result for providing examples and constructing families of SUM models by linking the univariate probability density functions (pdf's) $f_i(x_i)$, $i = 1,2$.

Lemma 2. *Let* $f_i(x_i)$, $i = 1, 2$ *be pdf's and* $\mathrm{g}(x_1, x_2)$ *a measurable function. Set*

$$f_\beta(x_1, x_2) = f_1(x_1)f_2(x_2) + \beta \mathrm{g}(x_1, x_2), \quad (x_1, x_2) \in \Re^2. \tag{6}$$

Then for some $\beta \in \Re, f_\beta(x_1, x_2)$ *is a SUM pdf with marginal pdf's* $f_i(x_i)$, $i = 1, 2$, *provided that:*

(a) $f_\beta(x_1, x_2) \geq 0$

(b) $\int_{\Re}\mathrm{g}(x_1, x_2)\mathrm{d}x_1 = \int_{\Re} \mathrm{g}\,(x_1, x_2)\mathrm{d}x_2 = 0$ *for all* $(x_1, x_2) \in \Re^2$

(c) $\int_{\Re} g(c-t,t)dt = 0$ *for all* $c \in \Re^2$

Proof. Condition (a) is required for $f_\beta(x_1, x_2)$ to be a pdf and (b) is needed for $f_i(x_i),\ i = 1,2$ to be marginal pdf's. Condition (c) is exactly what is needed to make $f_\beta(x_1, x_2)$ a SUM pdf. □

The next example illustrates Lemmas 1 and 2.

Example 1. Let $f_i(x_i),\ i = 1,2$ be two pdf's on [0, 1] and set

$$f_\beta(x_1, x_2) = f_1(x_1)f_2(x_2) + \beta \sin[2\pi(x_2 - x_1)],\ (x_1, x_2) \in [0,1]^2 \tag{7}$$

such that for some $\beta \in \Re, f_\beta(x_1, x_2)$ is a pdf on $[0, 1]^2$. Since $\sin[2\pi(x_2 - x_1)] = -\sin[2\pi(x_1 - x_2)]$, conditions of Lemma 2 are satisfied, and $f_\beta(x_1, x_2)$ is the pdf for a family of SUM models on the unit square. Two specific examples are as follows.

(a) Let $f_i(x_i) = 1,\ i = 1,2$ be the pdf of uniform distribution on [0, 1] and $\beta = -\frac{1}{2}$. Then, by Lemma 1, Spearman's rho (7) is $\rho_s = 0$. It can be shown that $\tau \neq 0$.

(b) Let $f_1(x_1) = \frac{1}{2} + x_1,\ f_2(x_2) = 1$, and $\beta = -\frac{1}{2}$. It can be shown that Kendall's tau and Spearman's rho for (7) are negative: $\tau = \frac{\pi - 4}{8\pi^3}$ and $\rho_s = \frac{-3}{4\pi^3}$.

We will develop more specific construction methods using $g(x_1, x_2) = f_1(x_1)f_2(x_2)q(x_1, x_2)$ in (6). We then have the pdf's in the following form:

$$f_\beta(x_1, x_2) = f_1(x_1)f_2(x_2)[1 + \beta q(x_1, x_2)], \quad (x_1, x_2) \in \Re^2, \tag{8}$$

where $f_i(x_i),\ i = 1,2$ are the marginal pdf's, $q(x_1, x_2)$ is a measurable bounded function on $\Re^2$ with bound $|q(x_1, x_2)| \leq B$, and $\beta = B^{-1}$. Various bivariate distributions in the form of (8) have been proposed in the literature, see, e.g., [6,8]. We will introduce two classes of SUM distributions in the form of (8).

The level of dependence in (8) is a function of β and the linking function $q(x_1, x_2)$. The following result facilitates calculation of the mutual information for the family (8).

Lemma 3. *The mutual information of bivariate distributions with pdf's of the form* (8) *is given by*

$$M_\beta(X_1, X_2) = \sum_{n=2}^{\infty} \frac{(-\beta)^n}{n(n-1)} E_2\{E_1[q(X_1, X_2)]^n\}, \tag{9}$$

where $E_i,\ i = 1,2$ *denotes the expectation with respect to* f_i.

Proof. Let

$$T(z) = (z+1)\log(z+1) = z + \sum_{n=2}^{\infty}(-1)^n \frac{z^n}{(n-1)n}, \tag{10}$$

where the second equality is the Taylor series expansion which converges uniformly for $|z| \le 1$. For $|\beta| = B^{-1}$, $|\beta q(x_1, x_2)|$, and we have

$$M_\beta(X_1, X_2) = \int\int_{\Re^2} f(x_1, x_2) \log \frac{f(x_1,x_2)}{f^*(x_1,x_2)} \mathrm{d}x_1 \mathrm{d}x_2$$

$$= \int\int_{\Re^2} f_1(x_1) f_2(x_2)\, T[\beta q(X_1, X_2)] \mathrm{d}x_1 \mathrm{d}x_2 \tag{11}$$

The result is obtained by applying (10) in (11), interchanging the integral and sum in (11), and noting that $E_2\{E_1[q(X_1, X_2)]\} = 0$, due to the normalization requirement. □

3. A Bivariate SUM Family

The following result presents a method for constructing a bivariate SUM family with given marginal distributions and gives the mutual information measure, Kendall's tau, and Spearman's rho for the family.

Proposition 1. *Let* $f_i(x) = f(x),\ i = 1, 2$ *in* (8) *be a symmetric pdf and the linking function* $q(x_1, x_2)$ *be such that*

$$-q(x_1, x_2) = q(x_2, x_1) = q(-x_1, x_2) = q(x_1, -x_2) \tag{12}$$

Then:

(a) *the bivariate function* (8) *is the pdf of a family of SUM distributions with marginals* $f_i(x_\iota) = f(x_\iota),\ i = 1, 2$, *and* $(a_1 X_1, a_2 X_2),\ a_i = \pm 1,\ i = 1, 2$ *are SUM variables;*

(b) *the mutual information for the family is given by*

$$M_\beta(X_1, X_2) = \sum_{n=1}^{\infty} \frac{\beta^{2n}}{(2n-1)2n} E_2\{E_1[q(X_1, X_2)]^{2n}\}, \tag{13}$$

where $E_i,\ i = 1, 2$ *denotes the expectation with respect to* f_i.

(c) *Kendall's tau and Spearman's rho are* $\tau = \rho_s = 0$.

Proof. It is easy to see that $f\beta(x1, x2)$ is a joint pdf.

(a) Let $g(x_1, x_2) = f_1(x_1) f_2(x_2) q(x_1, x_2)$. Then the first equality in (12) implies condition (c) and the second and third equalities in (12) imply condition (b) of Lemma 2. The proofs for distributions of $(a_1 X_1, a_2 X_2),\ a = \pm 1,\ i = 1, 2$ are similar.

(b) The mutual information is given by (9), where by the first equality in (12) the terms in the sum vanish for odd n, and we obtain (13).

(c) The pdf's and probability distribution functions of the family (8) are in the form of

$$f(x_1, x_2) = f_1(x_1)f_2(x_2) + \beta \mathrm{g}(x_1, x_2) \tag{14}$$

$$F(x_1, x_2) = F_1(x_1)F_2(x_2) + \beta G(x_1, x_2) \tag{15}$$

where $\mathrm{g}(x_1, x_2) = f_1(x_1)f_2(x_2)q(x_1, x_2)$ and $G(x_1, x_2) = \int_{-\infty}^{x_2}\int_{-\infty}^{x_1} g(t_1, t_2)\mathrm{d}t_1 t_2$.

Let I_{hk}, $h, k = 1, 2$ denote the integral of the product of the *h*th term in (14) and the *k*th term in (15). Clearly, $I_{11} = \frac{1}{4}$ and (4) and (5) for pdf's of the form (8) are given by

$$\tau = 4\beta(I_{12} + I_{21} + \beta I_{22}) \text{ and } p_s = 12\beta I_{21}. \tag{16}$$

Since $f_1(x_1)f_2(x_2) = f1(x2)f2(x1)$ and $G(x_1, x_2) = -G(x_2, x_1)$, the quantities in (16) are as follows.

$$I_{12} = \int\int_{\Re^2} G(x_1, x_2)f_1(x_1)f_2(x_2)\mathrm{d}x_1\,\mathrm{d}x_2 = 0.$$

Similarly, we obtain $I_{21} = 0$, which gives $\rho_s = 0$. We also have $\mathrm{g}(-x_1, x_2) = -\mathrm{g}(x_1, x_2)$ and $G(-x_1, x_2) = -G(x_1, x_2)$, so

$$I_{12} = \int\int_{\Re^2} G(x_1, x_2)\mathrm{g}(x_1, x_2)\mathrm{d}x_1\,\mathrm{d}x_2 = 0.$$

This is due to the fact that the inside integral is zero for every fixed x_2. Therefore $\tau = 0$. □

We see from (13) that $M_\beta(X_1, X_2)$ is an even and convex function of β. We can use partial sums of the sum on the right of (13) to approximate $M_\beta(X_1, X_2)$. For $\beta \approx 0$ we have

$$M_\beta(X_1, X_2) \approx \frac{1}{2}\beta^2 E_2\{E_1[q(X_1, X_2)]^2\}.$$

We can also bound the mutual information as

$$L_m \leq M_\beta(X_1, X_2) \leq U_m,$$

where

$$L_m = \sum_{n=1}^{m} \frac{\beta^{2n}}{(2n-1)2n} E_2\{E_1[q(X_1, X_2)]^{2n}\}, \quad m \geq 1,$$

and

$$U_m = \sum_{n=1}^{m} \alpha_n \beta^{2n} E_2\{E_1[q(X_1, X_2)]^{2n}\}, \tag{17}$$

$$\alpha_n = \begin{cases} \dfrac{1}{(2n-1)2n}, & \text{for } n < m, \\ \sum_{k=n}^{\infty} \dfrac{1}{(2k-1)2k}, & \text{for } n \geq m. \end{cases} \tag{18}$$

The lower bound for $M_\beta(X_1, X_2)$ is obtained by noting that the sum in (13) has nonnegative terms. The upper bound is obtained as follows. Since $|\beta q(x_1, x_2)| \leq 1$, if $k \geq m$ then

$$\beta^{2k}[q(x_1,x_2)]^{2k} \le \beta^{2m}[q(x_1,x_2)]^{2m}.$$

Therefore, for every $m \ge 1$,

$$M_\beta(X_1,X_2) \le \sum_{n=1}^{m} \alpha_n \beta^{2n} \int\int_{\Re^2} f_1(x_1)\, f_2(x_2)[q(x_1,x_2)]^{2n} \mathrm{d}x_1 \mathrm{d}x_2,$$

where α_n is defined in (18).

Proposition 1 is applicable in constructing SUM distributions by linking marginal distributions such as normal, Student t, and Laplace. The parameter β determines the strength of dependence and the linking function $q(x_1,x_2)$ determines the shape of the pdf. When $q(x_1,x_2)$ satisfies only the first equality, or if $f(x_i)$ is not symmetric, we still obtain a SUM distribution, but $f(x_i)$, $i = 1, 2$ are not the marginals anymore.

Next we provide two examples where the marginals are normal and the linking functions are the product of two functions

$$q(x_1,x_2) = C(x_1,x_2)K(x_1,x_2), \tag{19}$$

where $K(x_1,x_2)$ is the independent bivariate normal (BVN) kernel and $C(x_1,x_2)$ is specified in each example. More generally, *C*(*x*1, *x*2) can be any bivariate function such that

$$-C(x_1,x_2) = C(x_2,x_1) = C(-x_1,x_2) = C(x_1,-x_2),$$

and $K(x_1,x_2)$ can be the kernel of a circular bivariate distribution such as the bivariate Student t kernel $K_v(x_1,x_2) = (v + x_1^2 + x_2^2)^{-(v/2+1)}$, and the product of two Student t kernels $K_{v1,v2}(x_1,x_2) = \prod_{i=1}^{2}(v_i + x_i^2)^{-(v_i+1)/2}$.

Example 2. Let distributions of X_1 and X_2 be identical *N*(0, 1), and

$$q(x_1,x_2) = x_1x_2(x_1^2 - x_2^2)\mathrm{e}^{-\frac{1}{2}(x_1^2+x_2^2)}.$$

The upper bound B is obtained by changing to polar coordinates

$$|q(r\cos\theta, r\sin\theta)| = \left|\frac{1}{4}r^4 \sin(4\theta)\mathrm{e}^{-\frac{r^2}{2}}\right| \le B.$$

The maximum is at $r = 2$, from which we obtain $B = 4\mathrm{e}^{-2}$. The SUM model for (X_1,X_2) has pdf

$$f_\beta(x_1,x_2) = \frac{1}{2\pi}\mathrm{e}^{-\frac{1}{2}(x_1^2+x_2^2)}\left[1 + \beta x_1x_2(x_1^2 - x_2^2)\mathrm{e}^{-\frac{1}{2}(x_1^2+x_2^2)}\right], \quad (x_1,x_2) \in \Re^2,$$

where $0 \le \beta \le \frac{1}{4}\mathrm{e}^2 \approx 1.847$. The distribution of $S = X_1 + X_2$ is *N*(0, 2), given by the independent BVN model $f_0(x_1,x_2) = f(x_1)f(x_2)$.

The left side panels of Fig. 1 show the contour plots of the pdf's of this SUM family for $\beta = 0$ (independent BVN) and $\beta = 1$,1.847. These plots show patterns similar to that shown in Arnold and Strauss [15] for an interesting example where the model for the joint distribution was specified through normal conditionals; also see Arnold et al. [16] p. 69. These plots show that the densities are unimodal and as β increases the distribution becomes highly concentrated at the center. That is, the entropy of $f_\beta(x_1, x_2)$ is a decreasing function of β. Since the entropy of the marginal distribution does not depend on β, by (3), the mutual information increases with β. There is no closed form for (13), we use (17) to approximate its value for $\beta = \beta_0 \approx 1.847$ as:

$$M_\beta(X_1, X_2) \leq M_{\beta_0}(X_1, X_2) \approx 0.0959.$$

This bound is tight. The upper limit is equal to the mutual information of a BVN distribution with a correlation of approximately 0.42.

The regression function is

$$E[X_1 | X_2 = x_2] = \frac{\beta}{4\sqrt{2}} x_2(3 - 2x_2^2)\mathrm{e}^{-\frac{1}{2}x_2^2}.$$

Fig. 2(a) shows the plot of this highly nonlinear regression for $\beta = 1$, which reflects the uncorrelatedness between the two variables. The parameter β affects the amplitude, not the shape of the regression function.

Next we give an example where the SUM density is multimodal. We also obtain an explicit expression for the mutual information.

Example 3. Let distributions of X_1 and X_2 be identical *N*(0, 1), and

$$q(x_1, x_2) = \frac{x_1 x_2 (x_1^2 - x_2^2)}{(x_1^2 + x_2^2)^2} e^{\frac{1}{2}(x_1^2 + x_2^2)}.$$

The upper bound B is obtained by changing to polar coordinates

$$|q(r\cos\theta, r\sin\theta)| = \left|\frac{1}{4}\sin(4\theta)\, e^{-\frac{r^2}{2}}\right| \leq B,$$

which gives $B = \frac{1}{4}$ and $\beta \leq 4$. The SUM model for (X_1, X_2) has pdf

$$f_\beta(x_1, x_2) = \frac{1}{2\pi}\mathrm{e}^{-\frac{1}{2}(x_1^2 + x_2^2)}\left[1 + \beta\frac{x_1 x_2(x_1^2 - x_2^2)}{(x_1^2 - x_2^2)^2}\mathrm{e}^{-\frac{1}{2}(x_1^2 + x_2^2)}\right], \quad (x_1, x_2) \in \Re^2,$$

where $0 \leq \beta \leq 4$. The marginals are identical *N*(0, 1), so the distribution of $S = X_1 + X_2$ is *N*(0,2), given by the independent BVN model $f_0(x_1, x_2) = f(x_1)f(x_2)$.

The right side panels of Fig. 1 show the contour plots of the pdf's of this SUM family for β

= 1, 2, 4. These plots show that as β increases the distribution becomes highly concentrated at four modes. Thus, the entropy of $f_\beta(x_1, x_2)$ decreases and the mutual information increases with β. The mutual information is

$$M_\beta(X_1, X_2) = \log\left(1 + \sqrt{1 - \tfrac{\beta^2}{16}}\right) + \tfrac{2}{\beta}\arcsin\left(\tfrac{\beta}{4}\right) - \tfrac{1}{2}\sqrt{1 - \tfrac{\beta^2}{16}} - \log 2, \quad 0 \le \beta \le 4. \tag{20}$$

We find this expression directly by changing to polar coordinates:

$$M_\beta(X_1, X_2) = \frac{1}{2\pi}\int_0^\infty re^{-r^2/2} \int_0^{2\pi} [1 + h(r)\sin(4\theta)]\log[1 + h(r)\sin(4\theta)]\, \mathrm{d}\theta \mathrm{d}r,$$

where $h(r) = \frac{\beta}{4}\mathrm{e}^{-r^2/2}$. If $|u| \le 1$, then

$$\begin{aligned} S(u) &= \frac{1}{2\pi}\int_0^{2\pi} (1 + u\sin t)\log(1 + u\sin t)\mathrm{d}t \\ &= \log\left(\frac{1}{2}\sqrt{1 - u^2} + \frac{1}{2}\right) + 1 - \sqrt{1 - u^2}. \end{aligned}$$

Therefore,

$$M_\beta(X_1, X_2) = \int_0^\infty re^{\frac{-r^2}{2}} S\big(h(r)\big)\mathrm{d}r = \frac{4}{\beta}\int_0^{\beta/4} S(u)\mathrm{d}u.$$

This integral gives (20).

Since $M_\beta(X_1, X_2)$ is an increasing function of β,

$$M_\beta(X_1, X_2) \le M_4(X_1, X_2) = \tfrac{\pi}{4} - \log 2.$$

Note that $M_0 = 0$, which is the mutual information of the independent BVN limit and the upper limit is equal to the mutual information of a BVN distribution with a correlation of approximately 0.41.

The regression function is

$$E[X_1|X_2 = x_2] = \frac{\beta x_2}{\sqrt{2}}\left\{(1 + 2x_2^2)\mathrm{e}^{-\frac{1}{2}x_2^2} - 2\sqrt{\pi}|x_2|(x_2^2 + 1)[1 - \mathrm{erf}(|x_2|)]\mathrm{e}^{\frac{1}{2}x_2^2}\right\},$$

where $\mathrm{erf}(u) = \frac{2}{\sqrt{\pi}}\int_0^u e^{-z^2}\mathrm{d}z$ is the error function. Fig. 2(b) shows the plot of this highly nonlinear regression for $\beta = 1$, which reflects the uncorrelatedness between the two variables. Note that β affects the amplitude, not the shape of the regression function.

4. Comparisons

We compare the strength of dependence that can be captured by SUM models with models that do not possess SUM properties in three contexts: a discrete example, in a class of distributions that all powers of the two variables are uncorrelated, and with the bivariate F–G–M family of distributions.

The following example illustrates the SUM concept through a family of distributions constructed on a 3 × 3 grid which includes a SUM sub-family.

Example 4. Consider the bivariate family of distributions:

$$f_{\alpha,\beta}(x_1,x_2)=\begin{cases}\alpha, & \text{for } (0,0),(1,1),(2,2),\\ \beta, & \text{for } (0,1),(1,2),(2,0), \quad \alpha,\beta\geq 0, \alpha+\beta\leq\dfrac{1}{3},\\ \dfrac{1}{3}-(\alpha+\beta), & \text{for } (0,2),(1,0),(2,1),\end{cases}$$

The marginal distributions are uniform on $f_i(x_i)=\frac{1}{3},\ x_i=0,1,2$. It can be easily checked that for $\alpha=\frac{1}{9}$ the family $f_{\frac{1}{9},\beta}(x_1,x_2),\ 0\leq\beta\leq\frac{2}{9}$ is a SUM family, where the distribution of $S=X_1+X_2$ is given by the independent model $f_{\frac{1}{9},\frac{1}{9}}(x_1,x_2)=f(x_1)f(x_2)=\frac{1}{9}$. The mutual information function computed by (3) is

$$M_{\alpha,\beta}(X_1,X_2)=2\log 3+3\left[\alpha\log\alpha+\beta\log\beta+\left(\tfrac{1}{3}-\alpha-\beta\right)\log\left(\tfrac{1}{3}-\alpha-\beta\right)\right].$$

It can be shown that $M_{\alpha,\beta}(X_1,X_2)$ is convex in each parameter and for the SUM sub-family,

$$0=M_{\frac{1}{9},\frac{1}{9}}(X_1,X_2)\leq M_{\frac{1}{9},\beta}(X_1,X_2)\leq M_{\frac{1}{9},0}(X_1,X_2)=M_{\frac{1}{9},\frac{2}{9}}(X_1,X_2).$$

For a given β, $M_{\frac{1}{9},\beta}(X_1,X_2)$ can be more, less, or equal to $M_{\alpha.\beta}(X_1,X_2)$. That is, the dependence in the SUM sub-family can be stronger, weaker, or equal to that of a distribution which is not SUM. For example, $M_{\alpha,0}(X_1,X_2)=0.40, 0.46, 1.10$ for $\alpha=1/6, 1/9, 0$, respectively.

4.1. Bivariate SUM Models with Polynomial Dissociation

Consider distributions that have the following dissociation property:

$$\operatorname{cov}(X_1^n,X_2^m)=0,\quad \text{for all } m,n=1,2,\ldots$$

In this family all pairs of polynomial functions of the components are uncorrelated, thus we refer to (21) as polynomial dissociation.

Next we construct a family of SUM distributions with polynomial dissociation. We use the following result from Lukacs [17]. Let $\xi: \Re \rightarrow \Re$ be a function which is infinitely many times differentiable, vanishing outside [−0.5, 0.5], and $\int_{\Re} \xi^2(s)\mathrm{d}s = 1$. Then:

$$\psi(s) = \int_{\Re} \xi(u)\xi(s+u)\mathrm{d}u \tag{22}$$

is the characteristic function of a pdf $f(x)$, and $\Psi(s) = 0$ for $|s| \geq 1$ (Lukacs [17], Theorem 4.2.4).

Proposition 2. *Let* $f(x)$ *be the pdf with characteristic function* (22)*. Then*

(a) *the distributions with pdf's*

$$f_{j,k}(x_1, x_2) = f(x_1)f(x_2)\{1 + \cos[(2j+1)x_1]\cos[(2k+1)x_2]\}, \quad j \neq k = 0, 1, 2, \dots \tag{23}$$

are a family of SUM distributions with marginals $f_i(x_i) = f(x_i)$, $i = 1, 2$, *and* $\operatorname{cov}(X_1^n, X_2^m) = 0 \; for\; all\; m, n \;=\; 1, 2, \dots$;

(b) *the mutual information for the family is given by*

$$M(X_1, X_2) = \sum_{n=1}^{\infty} \frac{C_n^2}{2n(2n-1)}, \tag{24}$$

where $C_n = 2^{-2n}\binom{2n}{n}$.

Proof. (a) For *m* = 1, 2,... set

$$\psi_m(s) = \frac{1}{2}[\psi(s+m) + \psi(s-m)].$$

Then $\Psi_m(s) = 0$ unless $|s-m| < 1$ or $|s+m| < 1$, and $\Psi_m(s)$ is the Fourier transform of $\mathrm{g}_m(x) = f(x)\cos(mx)$.

Since the derivatives of $\Psi_m(s)$ at 0 all vanish, we get

$$\int_{\Re} x^n \mathrm{g}_m(x)\mathrm{d}x = 0\,, \quad \text{for } n = 0, 1, 2, \dots \tag{25}$$

Noting that $f(x_i)$, $i = 1, 2$ are pdf's, it immediately follows from (25) with $n = 0$ that $f_{j,k}(x_1, x_2)$ is a bivariate pdf. Now

$$\psi_{x_1,x_2}(s,t) = \psi(s)\psi(t) + \psi_{2j+1}(s)\psi_{2k+1}(t), \quad (s,t) \in \Re^2,$$

where $\psi_{2j+1}(t)\psi_{2k+1}(t) = 0$ for all $t \in \Re$, so $\psi_{X_1,X_2}(t,t) = [\psi(t)]^2$ for all $t \in \Re$. This shows that the distribution with pdf $f_{j,k}$ is a SUM distribution. Moreover, by (25), for every n,

$$\int_{\Re} x_i^n f_{j,k}(x_1, x_2)\mathrm{d}x_i = f(x_h)\int_{\Re} x_i^n f(x_i)\mathrm{d}x_i\,, \quad h \neq i = 1, 2,$$

and for every n and m,

$$E\left(X_i^n X_j^m\right) = E_1(X_1^n)E_2(X_2^m) + \int_{\mathfrak{R}} x_i^n \mathrm{g}_{2j+1}(x_1)\mathrm{d}x_1 \int_{\mathfrak{R}} x_2^m \mathrm{g}_{2k+1}(x_2)\mathrm{d}x_2$$
$$= E_1(X_1^n)E_2(X_2^m).$$

Thus, $\operatorname{cov}(X_1^n, X_2^m) = 0$.

(b) By Lemma 3, we have

$$M(X_1, X_2) = E_f\{\cos[(2j+1)X_1]\}E_f\{\cos[(2k+1)X_2]\}$$
$$+\sum_{n=2}^{\infty} \frac{(-1)^n}{n(n-1)} E_f\{\cos^n[(2j+1)X_1]\}E_f\{\cos^n[(2k+1)X_2]\}, \tag{26}$$

where E_f denotes the expected value with respect to the marginal pdf f. We express $\cos^n x$ in terms of $cos(hx)$, $h = 0,1,\ldots,n$, and use (25), we find that

$$E_f\{\cos^n[(2j+1)X_1]\} = E_f\{\cos^n[(2j+1)X_2]\} = \begin{cases} 0 & \text{if n is odd,} \\ C_{n/2} & \text{if } n \text{ is even.} \end{cases}$$

This completes the proof. □

The sum in (24) is of hypergeometric type but there appears to be no closed form expression for it. We can approximate it as $M(X_1, X_2) \approx 0.143329\ldots$ which corresponds to the mutual information of a bivariate normal distribution with a correlation of approximately 0.5.

A specific example of (23), $f_{0,1}$ was used in [18]. The SUM family (23) is in the class of bivariate distributions with pdf's

$$f_\beta(x_1, x_2) = f(x_1)f(x_2)[1 + \beta q_1(x_1)q_2(x_2)], \tag{27}$$

where $f(x)$ is a pdf and $q_i(x)$, $i = 1,2$ is periodic and bounded; see [19]. Alfonsi and Brigo [7] study copulas that are based on periodic functions. Next we show that (24) dominates the mutual information of another family of bivariate distributions with pdf's of the form (27) having the polynomial dissociation.

Consider the family of bivariate distributions with pdf's

$$f_\alpha(y_1, y_2) = f(y_1)f(y_2)[\quad], \quad y_1, y_2 > 0, \tag{28}$$

where

$$f(y) = \frac{1}{\sqrt{2\pi}y} \mathrm{e}^{-\frac{1}{2}(\log y)^2}$$

is the log-normal pdf and α is a positive parameter. It can be shown that (28) is a bivariate pdf with polynomial dissociation (21) but is not SUM. For $\alpha = 2\pi$, (28) gives the distribution used by De Paula [19]. We will show that

$$M_\alpha(Y_1, Y_2) < M(X_1, X_2) = \lim_{\alpha\to\infty} M_\alpha(Y_1, Y_2) = \sum_{n=1}^{\infty} \frac{C_n^2}{2n(2n-1)}. \tag{29}$$

That is, the SUM distribution (23) has stronger dependence than the non-SUM distribution (28). To show (29), let $V = logY$. By the invariance property of mutual information, $M(Y_1, Y_2) = M(V_1, V_2)$, where $V_i,\ i = 1,2$ are identically distributed variables as V having the standard normal distribution with pdf $h(v) = \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2}v^2}$. Letting $\beta q(v_1, v_2) = \sin(\alpha v_1)\sin(\alpha v_2)$ in (8), Lemma 3 gives

$$M(Y_1, Y_2) = \sum_{n=1}^{\infty} \frac{1}{2n(2n-1)} \left\{E_h\left[\sin^{2n}(\alpha V)\right]\right\}^2,$$

where E_h denotes the expectation with respect to $h(v)$. Using the trigonometric identity

$$\sin^{2n}(x) = C_n + 2^{1-2n} \sum_{k=1}^{n} (-1)^k \binom{2n}{n-k} \cos(2knx),$$

we have

$$E_h\left[\sin^{2n}(\alpha V)\right] = C_n + 2^{1-2n} \sum_{k=1}^{n} (-1)^k \binom{2n}{n-k} \mathrm{e}^{-2\alpha^2 k^2}.$$

It is easy to see that the sum of the terms with $k = 1$ and $k = 2$ is negative. Similarly, the sum of the terms with $k = 3,\ k = 4$ is negative and so on. Therefore, we obtain the inequality

$$E_h\left[\sin^{2n}(\alpha V)\right] < C_n.$$

Note that $\lim_{\alpha\to\infty} E_h\left[\sin^{2n}(\alpha V)\right] = C_n$. Therefore, we find that

$$M(Y_1, Y_2) < \sum_{n=1}^{\infty} \frac{C_n^2}{2n(2n-1)} \equiv M^*,$$

and $\lim_{\alpha\to\infty} M(Y_1, Y_2) = M^*$.

For $\alpha = 2\pi,\ M(X_1, X_2) - M(Y_1, Y_2) \approx 1.5 \times 10^{-33}$, so $M(Y_1, Y_2)$ is less than $M(X_1, X_2)$ but very close to $M(X_1, X_2)$.

4.2. Comparison with F–G–M Family

The pdf of distributions in the F–G–M family is in the form of

$$f(x_1,x_2) = f_1(x_1)f_2(x_2)\{1+\beta[1-2F_1(x_1)][1-2F_2(x_2)]\}, \quad |\beta| \le 1;$$

(see, e.g., [6], p. 114). Thus, the F–G–M distributions are in the family (8) with $q(x_1,x_2) = [1-2F_1(x_1)][1-2F_2(x_2)]$. The mutual information for the F–G–M bivariate family can be computed by Lemma 3. Noting that $U_i = F_i(X_i),\ i = 1,2$ have uniform distributions, we have

$$M_\beta(X_1,X_2) = M_\beta(U_1,U_2) = \textstyle\sum_{n=2}^{\infty} \frac{(-\beta)^n}{n(n-1)} E_1[(1-2U_1)^n]E_2[(1-2U_2)^n]. \tag{30}$$

Now

$$E_1[(1-2U_1)^n]E_2[(1-2U_2)^n] = E[(1-2U)^n]^2 = \frac{[1-(-1)^{n+1}]^2}{4(n+1)^2} = \begin{cases} 0 & \text{if } n \text{ odd} \\ \dfrac{1}{(n+1)^2} & \text{if } n \text{ even.} \end{cases}$$

Thus the terms in the sum (30) vanish for odd n, and we obtain

$$M_\beta(X_1,X_2) = \sum_{n=1}^{\infty} \frac{(-\beta)^{2n}}{2n(2n-1)(2n+1)^2}.$$

This confirms that dependence in the F–G–M family increases with $|\beta|$ and

$$M_\beta(X_1,X_2) = \sum_{n=1}^{\infty} \frac{(-\beta)^{2n}}{2n(2n-1)(2n+1)^2}.$$

where $M_\beta^*(X_1,X_2) = M_{|\beta|=1}(X_1,X_2)$. Computation using 10^6 terms indicates that $M^*(X_1,X_2) \approx 0.06$ and the series converges quickly; the first term in the sum is $\frac{1}{18} \approx 0.056$, the first 3 terms give 0.05957, and the first 10 terms give 0.05998. Thus, for the F–G–M family $M_\beta(X_1,X_2) \le 0.06$. However, the maximum strength of dependence for the F–G–M family $M^*(X_1,X_2) \approx 0.06$ is less than the maximum levels of dependence for the SUM distributions in Examples 2 and 3, $M_{\beta_0}(X_1,X_2) \approx 0.0959$ and $M_4(X_1,X_2) = \frac{\pi}{4} - \log 2 \approx 0.09225$, respectively. Interestingly, the Kendall's tau and Spearman's rho for the F–G–M family with $|\beta| = 1$ are $|\tau| = \frac{2}{9}$ and $|\rho_s| = \frac{1}{3}$ (see, e.g., [12]), but for distributions in Examples 2 and 3, $\tau = \rho_s = 0$. The maximum strength of dependence for the F–G–M family is also weaker than the dependence for the SUM family of Proposition 2, $M(X_1,X_2) \approx 0.1433$.

5. Multivariate SUM and POD (NOD)

Let F be the probability distribution function of $\mathbf{X} = (X_1, \dots, X_p)'$ and $\mathbf{X}^* = (X_1^*, \dots, X_p^*)'$ denote the random vector with probability distribution function $F^* = \prod_{i=1}^p F_i$, where F_i is the marginal probability distribution function of X_i.

Definition 2. F is said to be a SUM distribution of order $p(\mathrm{SUMp})$ if $\sum_{i=1}^p X_i \overset{st}{=} \sum_{i=1}^p X_i^*$.

Definition 2 can be extended to the product of a linear combination of marginals, that is $\mathbf{a}'\mathbf{X} \overset{st}{=} \mathbf{a}'\mathbf{X}^*$ where $\mathbf{a}' = (a_1, \dots, a_p)$. A particular case of interest is when $a_i = 0{,}1$, which leads to the following extension of Definition 2.

Definition 3. F is said to be a multivariate SUM distribution if it is SUMp and all n-dimensional marginal distributions, $n < p$ are SUMn. That is, $\mathbf{a}'\mathbf{X} \overset{st}{=} \mathbf{a}'\mathbf{X}^*$, for all $\mathbf{a}$'s such that $a_k = 0{,}1$ and $\sum_{k=1}^p a_k = n \leq p$.

The following examples show variants of SUM distributions.

Example 5. Let $\mathbf{X} = (X_1, X_2, X_3)'$.

(a) Consider the distribution with pdf

$$f_\beta(\mathbf{x}) = \frac{1}{(2\pi)^{\frac{3}{2}}} \mathrm{e}^{-\frac{1}{2}\mathbf{x}'\mathbf{x}} \left(1 + \beta(x_1 - x_2)(x_1 - x_3)(x_2 - x_3)\mathrm{e}^{-\frac{1}{2}\mathbf{x}'\mathbf{x}}\right), \quad \mathbf{x} \in \mathfrak{R}^3,$$

where $\beta = B^{-1}$ and

$$\left|(x_1 - x_2)(x_1 - x_3)(x_2 - x_3)\mathrm{e}^{-\frac{1}{2}\mathbf{x}'\mathbf{x}}\right| \leq B. \tag{31}$$

The characteristic function is

$$\psi_\beta(\mathbf{t}) = \mathrm{e}^{-\frac{1}{2}\mathbf{t}'\mathbf{t}} - \frac{1}{2^{9/2}}\beta i(t_1 - t_2)(t_1 - t_3)(t_2 - t_3)\mathrm{e}^{-\frac{1}{4}\mathbf{t}'\mathbf{t}}, \quad \mathbf{t} \in \mathfrak{R}^3,$$

where $\mathbf{t} = (t_1, t_2, t_3)'$. Clearly $f_\beta(\mathbf{x})$ is SUM3. It can be shown that $f_\beta(x_i, x_j)$, $i = j = 1, 2, 3$ are SUM2 for all β satisfying (31). So $f_\beta(\mathbf{x})$ is a trivariate SUM distribution. The univariate marginals are $N(0, 1)$, so the distribution of $S_n = \mathbf{a}'\mathbf{X}$ where $\sum_{k=1}^3 a_k = n \leq 3$ are $N(0, n)$, $n = 2, 3$ given by the independent trivariate normal model.

(b) Consider the distribution with pdf

$$f_\beta(\mathbf{x}) = \frac{1}{(2\pi)^{3/2}} \mathrm{e}^{-\frac{1}{2}\mathbf{x}'\mathbf{x}} \left[1 + \beta x_2(x_1^2 - x_3^2)\mathrm{e}^{-\frac{1}{2}\mathbf{x}'\mathbf{x}}\right], \quad \mathbf{x} \in \mathfrak{R}^3$$

where $\beta = B^{-1}$ and

$$\left| x_2(x_1^2 - x_3^2)\mathrm{e}^{-\frac{1}{2}\mathbf{x}'\mathbf{x}} \right| \leq B. \tag{32}$$

The characteristic function is

$$\psi_\beta(\mathbf{t}) = \mathrm{e}^{-\frac{1}{2}\mathbf{t}'\mathbf{t}} - \frac{1}{2^{9/2}}\beta i t_2(t_1^2 - t_3^2)\mathrm{e}^{-\frac{1}{4}\mathbf{t}'\mathbf{t}}, \quad \mathbf{t} \in \mathfrak{R}^3.$$

Clearly $f_\beta(\boldsymbol{x})$ is SUM3. It can be shown that for $\beta \neq 0$, $f_\beta(x_1, x_2)$ and $f_\beta(x_2, x_3)$ are not SUM2, and $f_\beta(x_1, x_3)$ is an independent BVN for all β satisfying (32). So $f_\beta(\mathbf{x})$ is SUM3, but not a trivariate SUM distribution. The univariate marginals are $N(0,1)$, so the distribution of $S_3 = X_1 + X_2 + X_3$ is $N(0,3)$, given by the independent trivariate normal model.

Example 6. Let $\mathbf{X} = (X_1, \ldots, X_p)'$ has pdf

$$f_\beta(\mathbf{x}) = \frac{1}{(2\pi)^{p/2}}\mathrm{e}^{-\frac{1}{2}\mathbf{x}'\mathbf{x}}\left[1 + \beta(x_1^2 - x_2^2)\mathrm{e}^{-\frac{1}{2}\mathbf{x}'\mathbf{x}}\prod_{k=1}^{p} x_k\right], \quad \mathbf{x} \in \mathfrak{R}^p$$

so that $\beta = B^{-1}$ and

$$\left|(x_1^2 - x_2^2)\mathrm{e}^{-\frac{1}{2}\mathbf{x}'\mathbf{x}}\prod_{k=1}^{p} x_k\right| \leq B.$$

The characteristic function is

$$\psi_\beta(\mathbf{t}) = \mathrm{e}^{-\frac{1}{2}\mathbf{t}'\mathbf{t}} - \frac{\beta}{4}\left(\frac{i}{2\sqrt{2}}\right)^{p/2}(t_1^2 - t_2^2)\mathrm{e}^{-\frac{1}{4}\mathbf{t}'\mathbf{t}}\sum_{k=1}^{p} t_k, \quad \mathbf{t} \in \mathfrak{R}^p,$$

where $\boldsymbol{t}' = (t_1, \ldots, t_p)$. Clearly $f_\beta(\mathbf{x})$ is SUMp. It can be shown that all n-dimensional marginals, $n < p$, are independent normal. So, $f_\beta(\mathbf{x})$ is a multivariate SUM distribution. The univariate marginals are $N(0,1)$, so the distribution of $S_n = \mathbf{a}'\mathbf{X}$ where $\sum_{k=1}^{p} a_k = n \leq p$ are $N(0,n)$, $n = 2, 3, \ldots, p$, given by the independent p-variate normal model.

Our final result relates the SUM distributions to the well-known notions of Positive Orthant Dependence (POD) and Negative Orthant Dependence (NOD) defined as follows.

Definition 4. A multivariate distribution *F* is said to be POD (NOD) if

$$\bar{F}(x_1, \dots, x_p) \geq (\leq) \sum_{i=1}^{p} \bar{F}_i(x_i),$$

where $\bar{F}(x_1, \dots, x_p) = \mathrm{pr}(X_1 > x_1, \dots, X_p > x_p)$ and $\bar{F}_i(x_i) = \mathrm{pr}(X_i > x_i)$.

It should be noted that POD (NOD) are the weakest among all existing notions of dependence. The special case of $p = 2$ is known as positive (negative) quadrant dependence. It is known that under POD (NOD), if $\rho(X_i, X_j) = 0$, the X_i and X_j are pairwise independent, without implying any higher order dependence among $(X_1, \dots, X_p)$. For details about POD (NOD) and other notions of dependence see Barlow and Proschan [20]. The following result shows that under POD (NOD), SUM models implies independence.

Lemma 4. *Let* **X** *be a nonnegative random vector with a POD (NOD) distribution F. Then F is a SUM distribution if and only if* $F(\mathbf{x}) = \prod_{i=1}^{p} F_i(x_i)$.

Proof. Independence implies SUM. We use induction to prove the converse for POD. For $n = 2$, POD implies $\bar{F}(x_1, x_2) \geq \bar{F}_1(x_1)\bar{F}_2(x_2)$. Since SUM implies uncorrelatedness,

$$\mathrm{cov}(X_1, X_2) = \int_0^\infty \int_0^\infty [\bar{F}(x_1, x_2) - \bar{F}_1(x_1)\bar{F}_2(x_2)]\, \mathrm{d}x_1 \mathrm{d}x_2 = 0.$$

Hence $\bar{F}(x_1, x_2) = \bar{F}_1(x_1)\bar{F}_2(x_2)$. Now suppose that the proposition holds for $r < p$. Using SUM property, $\Phi_{x_1 + \cdots + x_p}(t) = \Phi_1(t) \dots \Phi_p(t)$, where Φ denotes the moment generating function, and $X_i \geq 0$, $i = 1, \dots, p$ after some messy integrations by parts for any $m > r$ and, say for $t = -1$, we get

$$\begin{aligned} &\int_0^\infty \dots \int_0^\infty \mathrm{e}^{t(x_1 + \cdots + x_m)} \bar{F}(x_1, \dots, x_m)\, \mathrm{d}x_1 \dots \mathrm{d}x_m = \\ &\int_0^\infty \dots \int_0^\infty \mathrm{e}^{t(x_1 + \cdots + x_m)} \bar{F}_i(x_1) \dots \bar{F}_m(x_m) \mathrm{d}x_1 \dots \mathrm{d}x_m. \end{aligned} \tag{33}$$

For example, for $r = 2$, $m = 3$, and $t = -1$,

$$\begin{aligned} \Phi_{x_1+x_2+x_3} &= \int_0^\infty \int_0^\infty \int_0^\infty \mathrm{e}^{t(x_1+x_2+x_3)} f(x_1, x_2, x_3) \mathrm{d}x_1 \mathrm{d}x_2 \mathrm{d}x_3 \\ &= \Phi_1(t)\Phi_2(t) + t\int_0^\infty \mathrm{e}^{tx_3} \bar{F}_3(x_3) \mathrm{d}x_3 + t^2 \int_0^\infty \int_0^\infty \mathrm{e}^{t(x_1+x_3)} \bar{F}_{13}(x_1, x_3) \mathrm{d}x_1 \mathrm{d}x_3 \\ &\quad + t^2 \int_0^\infty \int_0^\infty \mathrm{e}^{t(x_2+x_3)} \bar{F}_{23}(x_2, x_3) \mathrm{d}x_2 \mathrm{d}x_3 \end{aligned}$$

$$+t^3 \int_0^\infty \int_0^\infty \int_0^\infty \mathrm{e}^{t(x_1+x_2+x_3)} \bar{F}(x_1, x_2, x_3) \mathrm{d}x_1 \mathrm{d}x_2 \mathrm{d}x_3,$$

where $\bar{F}_{ij}(x_i, x_j)$ is the bivariate survival function of (X_i, X_j). Similarly, $\Phi_{X_1}(t)\Phi_{X_2}(t)\Phi_{X_3}(t)$ is given by the same expression as above where $\bar{F}(x_1, x_2, x_3)$ in the last integral is replaced with $\bar{F}_1(x_1)\bar{F}_2(x_2)\bar{F}_3(x_3)$.

From (33) we have

$$\int_0^\infty \dots \int_0^\infty \mathrm{e}^{t(x_1+\dots+x_m)} [\bar{F}(x_1, \dots, x_m) - \bar{F}_1(x_1) \dots \bar{F}_m(x_m)] \, \mathrm{d}x_1 \dots \mathrm{d}x_m = 0.$$

Since F is POD, the integrand is nonnegative and the equality is attained if and only if $\bar{F}(x_1, \dots, x_m) = \bar{F}_1(x_1) \dots \bar{F}_m(x_m)$ for all $\mathbf{x}$, i.e., $X_1, \dots, X_m$ are independent. Proof for NOD is similar. □

6. Conclusions

The SUM distributions can provide solution for some modeling applications where the variable of interest consists of the sum of a few components. Examples include household income, the total profit of major firms in an industry, and a regression model $Y = \mathrm{g}(X) + \epsilon$ where $\mathrm{g}(X)$ and ϵ are uncorrelated (the standard assumption), however, they may not be independent. For example, in Bazargan et al. [21], the return value of significant wave height (Y) is modeled by the sum of a cyclic function of random time delay $\hat{\mathrm{g}}(D)$ and a residual term $\hat{\epsilon}$. They found that the two components are uncorrelated but not independent and used (1) to calculate the distribution of the return value.

We showed how to construct bivariate SUM models for applications. At a general level, the product marginal pdf's of marginals are added to a multiple of a bivariate function $\mathrm{g}(x_1, x_2)$ which integrates to zero and changes sign when we interchange x_1 with x_2. Another construction produces bivariate SUM models with identical symmetric marginal distributions such as normal, Student t, and Laplace. In practice, one may rather easily develop models for the univariate distributions of each component and test for independence and lack of correlation between them. If tests reject independence but not lack of correlation, a SUM model can be appropriate. The linking function $q(x_1, x_2)$ models the dependence and determines the shape of the regression function. Selection of $q(x_1, x_2)$ can be a challenging task. We provided two examples for linking normal marginal distributions into SUM models.

We showed that Kendall's tau and Spearman's rho can fail for measuring dependence between SUM variables. We developed formulas for the mutual information measures that enabled us to assess the strengths of dependence captured by examples of SUM distributions and to make comparison with models that do not possess SUM properties. Using a discrete example, we showed that the strength of dependence in a SUM sub-family can be stronger, weaker, or equal to that of other distributions in the family which are not SUM. We also showed that the SUM models are capable of capturing higher levels of dependence than the maximum strength of dependence for the F–G–M family. Finally, we proved that in the class of POD (NOD) distributions, the SUM model implies independence, so for these classes the product of marginals cannot be used for computing the distribution of the sum without independence. Fitting SUM models to the data and simulating from SUM distributions are topics of future research.

References

[1] G.G. Hamedani, M.N. Tata, On the determination of the bivariate normal distribution from distributions of linear combinations of the variables, American Mathematical Monthly 82 (1975) 913–915.

[2] T.M. Durairajan, A classroom note on sub-independence, Gujarat Statistical Review VI (1979) 17–18.

[3] G.G. Hamedani, G.G. Walter, A fixed point theorem and its application to the Central Limit Theorem, Archiv Der Mathematik 43 (1984) 258–264.

[4] G.G. Hamedani, E.S. Key, H.W. Volkmer, Solution to a functional equation and its application to stable and stable-type distributions, Statistics and Probability Letters 69 (2004) 1–9.

[5] T.M. Cover, J.A. Thomas, Elements of Information Theory, Wiley, New York, 1991.

[6] D. Drouet-Mari, S. Kotz, Correlation and Dependence, Imperial College Press, London, 2001.

[7] A. Alfonsi, D. Brigo, New families of Copulas based on periodic functions, Communications in Statistics-Theory and Methods 34 (2005) 1437–1447.

[8] C. Balakrishnan, C. Lai, Continuous Bivariate Distributions, second ed., Springer, New York, 2009.

[9] M. Fischer, I. Klein, Constructing generalized FGM copulas by means of certain univariate distributions, Metrika 65 (2006) 243–260.

[10] M.C. Jones, A dependent bivariate *t* distribution with marginals on different degrees of freedom, Statistics and Probability Letters 56 (2002) 163–170.

[11] M.C. Jones, Changing marginals of multivariate densities by multiplication, with application to skewing spherically symmetric distributions, Journal of Multivariate

Analysis 81 (2002) 85–99.

[12] R.B. Nelsen, An Introduction to Copulas, second ed., Springer, New York, 2006.

[13] J.A. Rodríguez-Lallena, M. Úbeda-Flores, A new class of bivariate copulas, Statistics and Probability Letters 66 (2004) 315–325.

[14] W.T. Shaw, K.T.A. Lee, Bivariate Student t distributions with variable marginal degrees of freedom and independence, Journal of Multivariate Analysis 99 (2008) 1276–1287.

[15] B.C. Arnold, D. Strauss, Bivariate distributions with conditionals in prescribed exponential families, Journal of the Royal Statistical Society Series B 53 (1991) 365–375.

[16] B.C. Arnold, E. Castillo, J.M. Sarabia, Conditional Specification of Statistical Models, Springer, New York, 1999.

[17] E. Lukacs, Characteristic Functions, second ed., Griffin, London, 1970.

[18] G.G. Hamedani, H.W. Volkmer, Letter to the Editor, The American Statistician 63 (2009) 295.

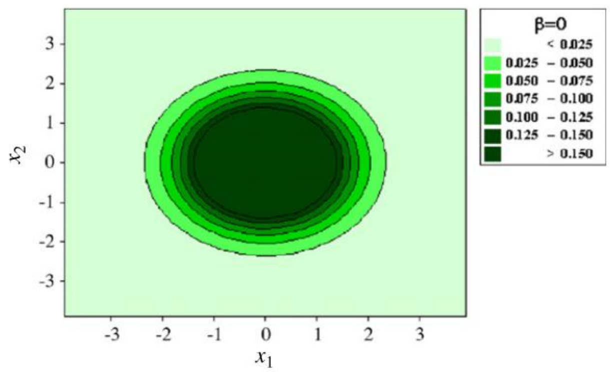
[19] A. De Paula, Conditional moments and independence, The American Statistician 62 (2008) 219–221.

[20] R.E. Barlow, F. Proschan, Statistical Theory of Reliability and Life Testing: Probability Models, To Begin with, Silver Spring, MD, 1981.

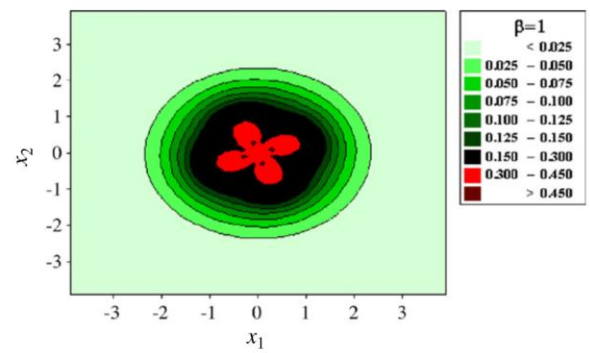
[21] H. Bazargan, H. Bahai, A. Aminzadeh-Gohari, Calculating the return value using a mathematical model of significant wave height, Journal of Marine Science and Technology 12 (2007) 34–42.

Appendix

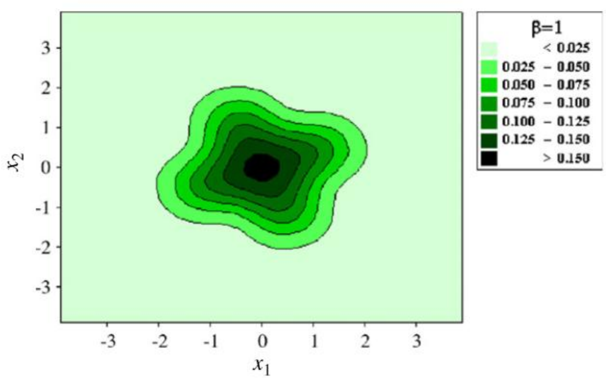
Figure 1: Contour Plots of SUM Models in Examples 2 and 3



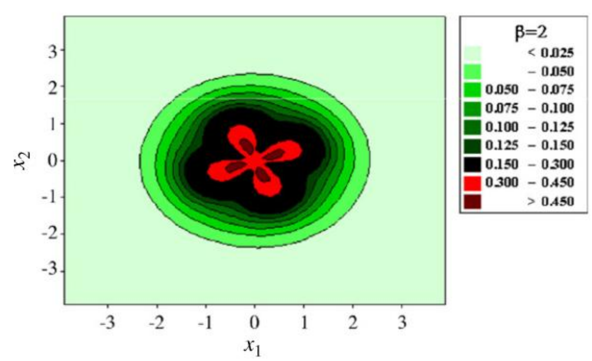
(a) Bivariate normal.



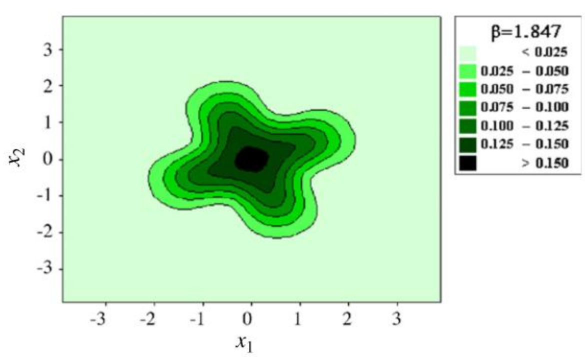
(d) Example 3.



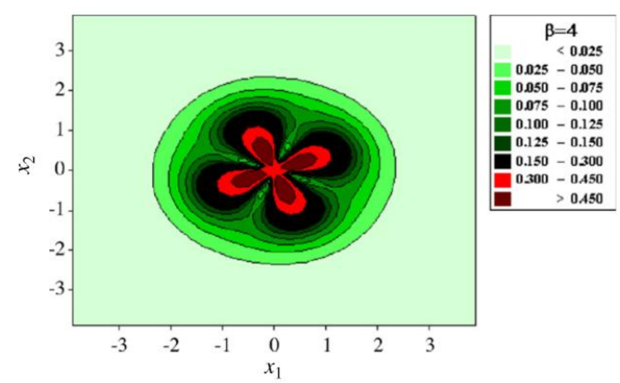
(b) Example 2.



(e) Example 3.

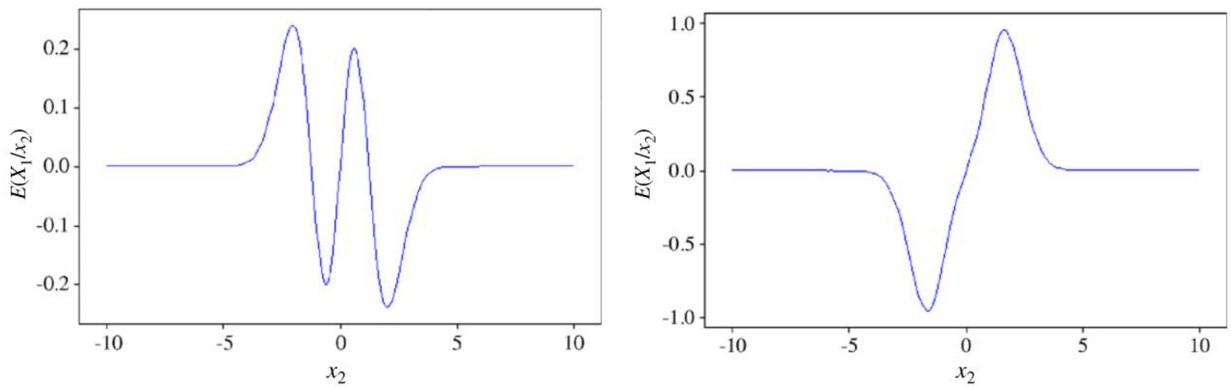


(c) Example 2.



(f) Example 3.

Figure 2: Regression Plots of Two SUM Models in Examples 2 and 3 with $\beta = 1$



(a) Regression function for model in Example 2.

(b) Regression function for model in Example 3.